(72) Inventors:
• Yamashita, Kei,
Nippon Telegraph & Tel. Corp
Tokyo 163-1419 (JP)
• Tanaka, Hiroyuki,
Nippon Telegraph & Tel. Corp
Tokyo 163-1419 (JP)
• Utsumi, Shusuke,
c/o Sony Corporation
Tokyo (JP)
• Shionozaki, Atsushi,
Sony Comp. Science Lab. Inc.
Tokyo 141 (JP)

(74) Representative:
Melzer, Wolfgang, Dipl.-Ing. et al
Patentanwälte
Mitscherlich & Partner,
Sonnenstrasse 33
80331 München (DE)

(54) **Communication method and data communications terminal with data communication protocol for inter-layer flow control**

(57)   A data communication protocol is implemented to control data flow between a transport layer (13) and a data link layer (11). The protocol provides two manners of data flow control. In the first control manner, when the occupancy of packet buffers of the data link layer (11) is smaller than a first threshold, the data link layer informs of the transport layer that the packet transport should be suspended, and thereafter when the occupancy is smaller than a second threshold smaller than the first one, the data link layer requests the transport layer (13) to resume the packet transfer. In the second control manner, when the amount of packets is smaller than a value determined based on the transfer rate and maximum burst size, the transport layer transfers a packet to a lower layer, and when the packet amount is larger than the value, a packet is transferred after a predetermined time passes to control the transfer rate to a predetermined one.
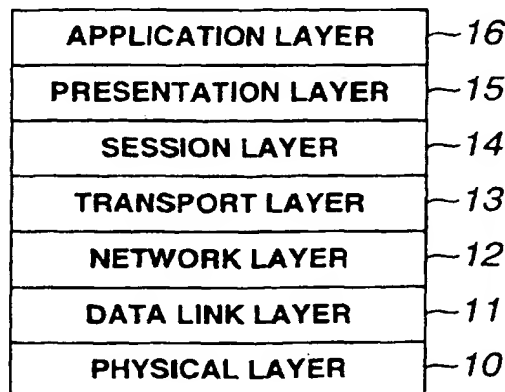
APPLICATION LAYER	—16
PRESENTATION LAYER	—15
SESSION LAYER	—14
TRANSPORT LAYER	—13
NETWORK LAYER	—12
DATA LINK LAYER	—11
PHYSICAL LAYER	—10

FIG.1

EP 0 905 950 A1

Description

BACKGROUND OF THE INVENTION

1. Field of the Invention:

[0001] The present invention relates to a communication method with a data communication protocol for packet transfer a transport layer and a data link layer, and a data communications terminal using the data communication protocol for data communication.

2. Description of Related Art:

[0002] The data communication protocol being the interfacing condition for data communications, typically the TCP/IP (Transmission Control Protocol/Internet Protocol) and OSI (Open Systems Interconnection) protocols, has a hierarchical structure as shown in FIG. 1. As shown, the hierarchical structure consists, at the level of operations executable by application programs, of an application layer 16 containing characteristics of data communications terminals, coding control function, etc., presentation layer 15 containing communication-oriented profiles, document-oriented profiles, etc., session layer 14 having a control function for data transmission and reception IDs, etc., transport layer 13 for carrying out a transport procedure independent of any network because the transport layer 13 is an upper layer, network layer 12 for control of call setting, disconnection, etc. for each network, data link layer 11 for establishing, releasing, error control, etc. of a data link, and a physical layer 10 for setting a user and network voltage physical interface conditions.

[0003] For packet transmission, a packet generated by a software of the application layer 16 (also including the presentation layer 15 and session 14) is transferred to a software which implements the transport layer 13, and then sequentially from the transport layer 13 to the network layer 12, from the network layer 12 to the data link layer 11, and further to a lower layer. Finally, the packet is represented in the form of an electric or light pulse for transmission from a data communications terminal.

[0004] In the conventional packaged data communications terminal, the packet is copied from a memory only twice. That is, the memory copying is done when the packet is moved from the application layer 16 to the transport layer 13 in a main memory 1 and when the packet is moved from the main memory 1 of the data communications terminal to a packet buffer 3 in a network interface module 11 via a bus 5, as shown in FIG. 2. The packet is moved between other layers higher than the data link layer 11 just by passing a pointer to a location in the main memory 1 where the packet is stored.

[0005] In the memory copying from the application layer 16 to the transport layer 13, a flow control is made to avoid any data loss due to a packet buffer overflow in the transport layer 13. Namely, when the packet buffer in the transport layer 13 is about to overflow, such a control will be done that the application layer 16 suspends packet copying. However, data flow between the data link layer 11 and a higher layer will not be controlled. The packet buffers of the data link layer 11 include one in the network interface module 2 and one in the main memory 1. Since no flow control is done, however, even if both the buffers overflow, the higher layer will continuously transfer packets to the data link layer 11 with the result that the packets will be discarded.

[0006] The reason for the above lies in the basic design policy for the conventional data communications protocol that the reliability of data communications is not assured for the layers lower than the transport layer 13 and the interaction between the layers is minimized.

[0007] The reliability of the data communications is assured only by packet retransmission from the transport layer 13. Therefore, even if there is no space for packets in the packet buffers of the data link layer 11, the protocol in the transport layer 13 does not suspend the packet transmission, thus the packets are lost and thereafter packets are retransmitted.

[0008] Conventionally, however, a packet is generated in the application layer 16, and transferred to the date link layer 11 via the transport layer 13 and network layer 12 at a lower rate than the rate of packet transmission from the network interface module 2 to outside the terminal, so that there occurs little packet loss due to overflow of the packet buffer in the data link layer 11.

[0009] The conventional data communication protocol is likely to be much deteriorated in throughput of data communications as the terminal performance and network function have become higher. More particularly, since the performance of the terminal components such as CPU, etc. has rapidly become improved, the rate of packet transfer to the data link 11 from a higher layer is higher than that of the packet transmission from the network interface module 2 to outside the terminal in some cases. In such a case, there is a possibility that the packet buffers of the data link layer 11 overflow, resulting in a packet loss even in the terminal.

[0010] Also, the data link layer 11 has functionally been improved to freely change the packet transmission rate as in the ATM (asynchronous transmission mode). In this case, there is a possibility that the rate of packet transmission from the data link layer 11 to outside the terminal is controlled to a low one depending upon how much congested the network is. As the packet transmission rate is thus low, the difference between the low rate of packet transmission from the data link layer 11 to outside the terminal and the rate of packet transfer to the data link layer 11 from a higher layer is larger, the possibility of the packet loss will be larger.

[0011] When a packet is lost, the transport layer 13 has to retransmit a packet, so that the data transmission throughput will be lower and the bandwidth of the network be wasted. Particularly in a transport protocol for

congestion control such as TCP, when a packet loss is detected, the packet transmission rate is autonomously controlled to a low one by reducing the transmission window size or the like, the data transmission throughput will be remarkably low.

[0012]   FIG. 3 shows the characteristic curve showing TCP throughput vs. maximum transmission window size.

[0013]   The round trip time (RTT) between a transmitting terminal and a receiving terminal is 20 ms. With a large RTT, the transmission window size has to be increased in order to provide a higher data transmission throughput by TCP. The transmission window size referred to herein means a maximum data length that can be batch-transmitted with response of no reception acknowledgment. If the transmission window size is increased excessively, the packet buffers of the data link layer 11 overflow, resulting in a packet loss. This is caused by the fact that a higher layer will batch-transfer data in an amount for the transmission window size to the data link layer 11. If a packet is lost, the availability control of TCP is put into action to largely lower the data transmission throughput. The conventional data communication protocol is disadvantageous in that the to-be-avoided packet loss inside the terminal lowers the data transmission throughput and network availability.

SUMMARY OF THE INVENTION

[0014]   Accordingly, the present invention has an object to overcome the above-mentioned drawbacks of the prior art by providing a communication method and data communications terminal in which a packet transferred between a transport layer and a data link layer is not discarded and thus a packet is lost inside the terminal, thereby preventing the data transmission throughout and network availability from being lower.

[0015]   The above object can be accomplished by providing a communication method in which data flow between a transport layer and a data link layer is controlled. That is, a packet transfer from the transport layer is controlled to prevent packet buffers of the data link layer from overflowing by any of the following two methods:

[0016]   In the first method of flow control, when the packet buffers of the data link layer is so highly occupied that the buffers are likely to overflow, it is informed to the transport layer that the data link layer should suspend the packet transfer. In the second manner, a packet transmission rate of the data link layer is informed to the transport layer which in turn will transfer the packet to a lower layer at a rate lower than the informed one.

[0017]   According to the present invention, since the flow control is done between the transport and data link layers, no packet loss takes place inside the terminal and the data transmission throughput and network availability can be prevented from being lower.

BRIEF DESCRIPTION OF THE DRAWINGS

[0018]   These objects and other objects, features and advantages of the present intention will become more apparent from the following detailed description of the present invention when taken in conjunction with the accompanying drawings, of which:

FIG. 1 shows a hierarchical structure of the data communication protocol;
FIG. 2 shows a configuration of the data communications terminal according to the present invention;
FIG. 3 is a characteristic curve of the data transmission throughput attainable by the present invention and that by the conventional TCP;
FIG. 4 is a functional flow chart of the flow control according to a first embodiment of the present invention; and
FIG. 5 is a flow chart of a traffic shaping according to a second embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0019]   FIG. 2 shows the configuration of the data communications terminal according to the present invention, and FIG. 3 shows the data transmission throughput attainable when the flow control according to the present invention is adopted.

[0020]   As seen from FIG. 2, a communication protocol is implemented by a software and executed by a central processing unit (CPU) 4. A packet to be transmitted is generated in a main memory 1 by an application layer. The packet is moved into a bucket buffer 3 in a network interface module 2 by a software which implements a data link layer. The packet is transmitted to outside from the terminal by the network interface module 2.

[0021]   FIG. 3 shows a TCP with flow control according to the present invention and a conventional TCP for comparison in data transmission throughput with each other. In the TCP window size control, data transmission and reception are controlled depending upon a number of packets (window size) continuously receivable correspondingly to the size of the buffer in a receiving terminal. In the conventional TCP, when the maximum window size exceeds about 170 kB, the data transmission throughput is lower. On the contrary, the TCP with flow control according to the present invention permits to increase the data transmission throughput with no lowering thereof. In the TCP with flow control, a function to keep the rate of data transfer from the transport layer at a designated one is added to a conventional TCP. The designated rate is generally same as a actual data transmission throughput as graphically shown in FIG. 3. Different from the conventional TCP, the TCP with flow control according to the present invention is such that even if the transmission window size is large, the data transmission throughput will not

be lower, as seen from FIG. 3. This is because no packet will be lost.

First embodiment:

[0022] FIG. 4 is a functional flow chart of the flow control according to the second embodiment of the present invention.

[0023] The first embodiment adopts the first flow control method in which when the packet buffers of the data link layer are so highly occupied that a bucket packet is likely to overflow, it is informed to the transport layer that the data link layer should suspend the packet transfer.

[0024] In this first method of flow control, the main memory 1 has prepared therein a variable indicative of a probability of packet buffer overflow (will be referred to as "buffer overflow flag" hereinafter) which can be accessed from both the software implementing the data link layer (will be referred to as "data link layer program" hereinafter) and the software implementing the transport layer (will be referred to as "transport layer program" hereinafter).

[0025] FIG. 4A shows the operations of the data link layer program, and FIG. 4B shows the operations of the transport layer program. When the occupancy or activity ratio of the bucket buffers of the data link layer is below a threshold T1, the data link layer program will set or raise a buffer overflow flag (at Step 201). On the other hand, the transport layer program checks the buffer overflow flag before transferring a packet to a lower layer (at Step 211) to discriminate whether or not the flag is set and suspend the packet transmission if the flag is set (at Step 212). Thereafter, when the occupancy of the packet buffers of the data link layer is below a threshold T2, the data link layer program lowers the buffer overflow flag (at Steps 203 and 204), revokes and requests the transport layer program to resume the packet transmission (at Step 205). It should be noted that the thresholds T1 and T2 are a value (T1 > T2) depending upon the characteristic of the network interface module 2. When knowing the arrival from the data link layer of the request to resume the packet transmission (at Step 214), the transport layer program will resume the packet transmission (at Step 215). Then, the transport layer will transmit packets (at Steps 211 and 213) while making sure that no buffer overflow flag is set for the data link layer, until there exists no further packet to be transmitted (at Step 216).

Second embodiment:

[0026] The second embodiment of the present invention employs the second method of flow control in which a packet transmission rate of the data link layer is informed to the transport layer which in turn will transfer the packet to a lower layer at a rate lower than the informed one.

[0027] In this second method of flow control, when a data transmission is started, the data link layer program will first inform the transport layer program of a rate of data transmission from the data link layer. Also, after the data transmission is started, each time the rate of data transmission from the data link layer changes, the data link layer program will inform the transport layer of the new data transmission rate. The transport layer program has provided therein a traffic shaping function, that is to say, a function to control the rate of data transmission to a lower layer, so that packet transfer is done at a rate lower than the data transmission rate of the data link layer.

[0028] More particularly, to control the rate of packet transfer from the transport layer to the data link layer to a predetermined value Rt lower than the rate of packet transfer from the data link layer, the transport layer transfers a packet to a lower layer at a time only when a total Psum + P of a sum Psum of data for packets having been transmitted for a period from a time t - s until t and a size P of a packet about to be transmitted is smaller than a value $Rt \times (t - s) + Bmax$ which is determined based on a packet transfer rate Rt and maximum burst size Bmax at the time t - s after the first transfer of packet. In other cases, no packet transfer is done before a time t' after the time t at which the above conditions are met, and at the time t', a packet transfer is done, thereby attaining the above object.

[0029] FIG. 5 is a flow chart of a traffic shaping according to the second embodiment of the present invention.

[0030] In the embodiment shown in FIG. 5, the second method of flow control is effected for an arbitrary period S_0 from the packet transmission start time t - s until the current time t in the expression for determination of when packet transmission is done.

[0031] The transport layer program has two parameters: Data transfer rate Rt and maximum burst size Bmax, and a token T being a variable. The initial value of the token is equal to the maximum burst size Bmax. Conceptually, the token increases at the data transfer rate Rt as the time passes but not above the maximum burst size Bmax.

[0032] After idling (at Step 100), when a packet arrives from the application layer (at Step 101), first the token T is renewed (at Step 103). Namely, the token T is renewed to a value $T = min(Bmax, T + Rt \times (t - tp))$ where t is the current time and tp is the preceding time of packet transmission.

[0033] Next, the packet size P ia compared with the token T (at Step 104). When P < T, a packet is transmitted (at Step 106). When the packet is transmitted, the token T is decreased by the packet size P ($T = T - P$). After the packet transmission, it is judged whether there exists a packet to transmit (at Step 107). If there is no further packet to be transmitted, the operation returns to the idling status (at Step 100). When there remains a packet or packets to transmit, the packet transmission is processed for time-out (at Step 102).

[0034] On the other hand, if the result of comparison shows that the token T is smaller than P, no packet is transmitted. A time at which the token T is larger than P is calculated (to = t + (P - T)/Rt) where to is a time-out time to set a time length until the time-out operation (at Step 105).

[0035] When the time-out operation is done, a time-out processing starts (at Step 102), the transport layer program is revoked and the packet transfer is resumed (at Step 101).

[0036] It should be noted that the data transfer rate Rt has to be lower than the data transmission rate R1 of the data link layer that is informed from the data link layer. Also, note that the maximum burst size Bmax should be smaller than Buf × (Rmax - R1)/Rmax where Buf is a packet buffer size in the data link layer and Rmax is a maximum rate of transfer from the transport layer to a lower layer to avoid a buffer overflow at the data link layer.

[0037] As having been described in the foregoing, the present invention provides a flow control between the transport layer and data link layer so that no packet is lost in the terminal, whereby it is possible to prevent the data transmission throughput and network availability from being lower.

Claims

1. A communication method with a data communication protocol having a hierarchical structure consisting of, from above, an application layer (16), presentation layer (15), session layer (14), transport layer (13), network layer (12), data link layer (11) and a physical layer (10), said data communication protocol being such that data flow between the transport and data link layers is controlled by informing the transport layer (13) that the data link layer (11) should suspend a packet transfer from the transport layer (13) to the data link layer (11) via the network layer (12) when the bucket buffer (3) in the data link is likely to overflow, or by informing the transport layer of a packet transmission rate of the data link layer and transferring the packet from the transport layer at a rate lower than the informed packet transmission rate.

2. The method as set forth in Claim 1, wherein the flow control is done through information by the data link layer of the transport layer of an instruction to suspend the packet transfer when the availability of the bucket buffer in the data link layer is larger than a first threshold, and then through request by the data link layer to he transport layer to resume the packet transferred when the availability of the packet buffer is smaller than a second threshold smaller than the first threshold.

3. The method as set forth in Claim 1 or 2, wherein the

flow control is done by transferring from the transport layer a packet to a lower layer only when a total Psum + P of a sum Psum of data for packets having been transmitted for a period from a time t - s until t and a size P of a packet about to be transmitted is smaller than a value Rt (t - s) + Bmax which is determined based on a packet transfer rate Rt and maximum burst size Bmax at the time t - s after the first transfer of packet, but by transferring, in other cases than above, no packet transfer before a time t' after the time t at which the above conditions are met, and a packet at the time t', to control the rate of packet transfer from the transport layer to the data link layer to a predetermined value Rt smaller than the rate of packet transfer from the data link layer.

4. A data communications terminal having a central processing unit (4), main memory (1) and a network interface module (2) connected via a bus (5) between the central processing (4) unit and main memory (1), the flow control as set forth in any of Claims 1 to 3 being done by the central processing unit (4).
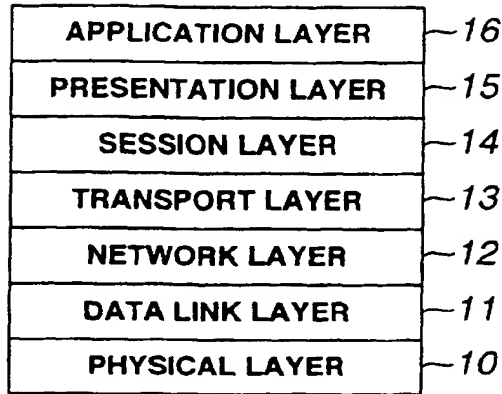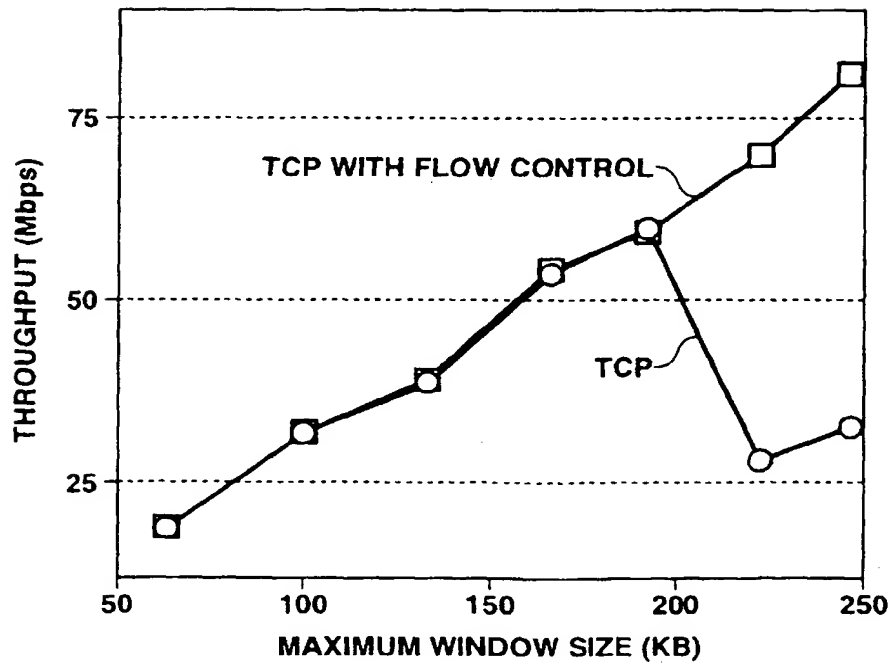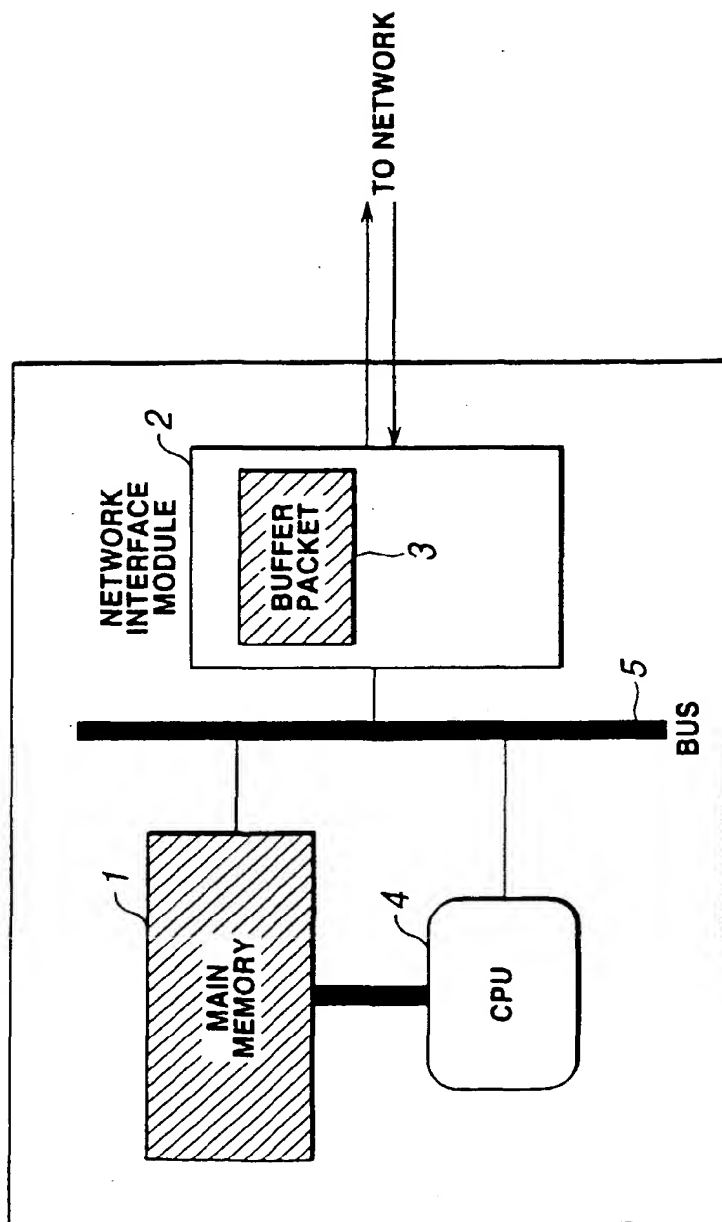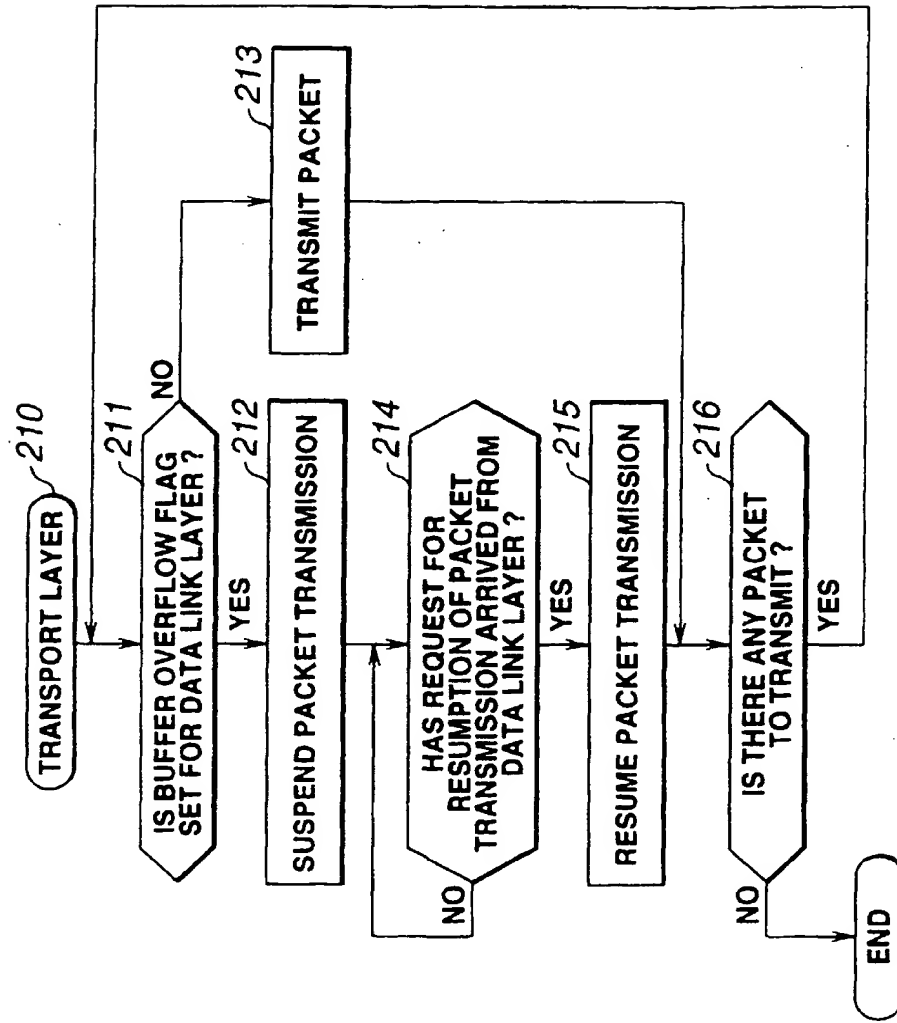
APPLICATION LAYER	~16
PRESENTATION LAYER	~15
SESSION LAYER	~14
TRANSPORT LAYER	~13
NETWORK LAYER	~12
DATA LINK LAYER	~11
PHYSICAL LAYER	~10

FIG.1

FIG.3

FIG.2

FIG.4B

FIG.4A

8

FIG.5

IDLING — 100

PACKET ARRIVES FROM APPLICATION LAYER — 101

TIME-OUT PROCESSING — 102

TOKEN T IS RENEWED
$T = \min(Bmax, T + Rt \times (t - tp))$ — 103

$T > P$? — 104

TIME-OUT IS SET
$to = t + (P - T) / Rt$ — 105

PACKET IS TRANSMITTED
$T = T - P$ — 106

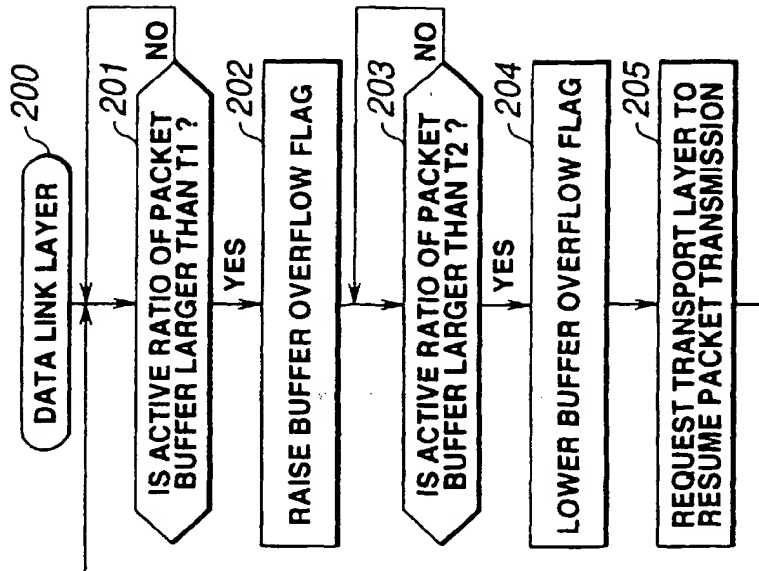IS THERE ANY PACKET TO TRANSMIT ? — 107
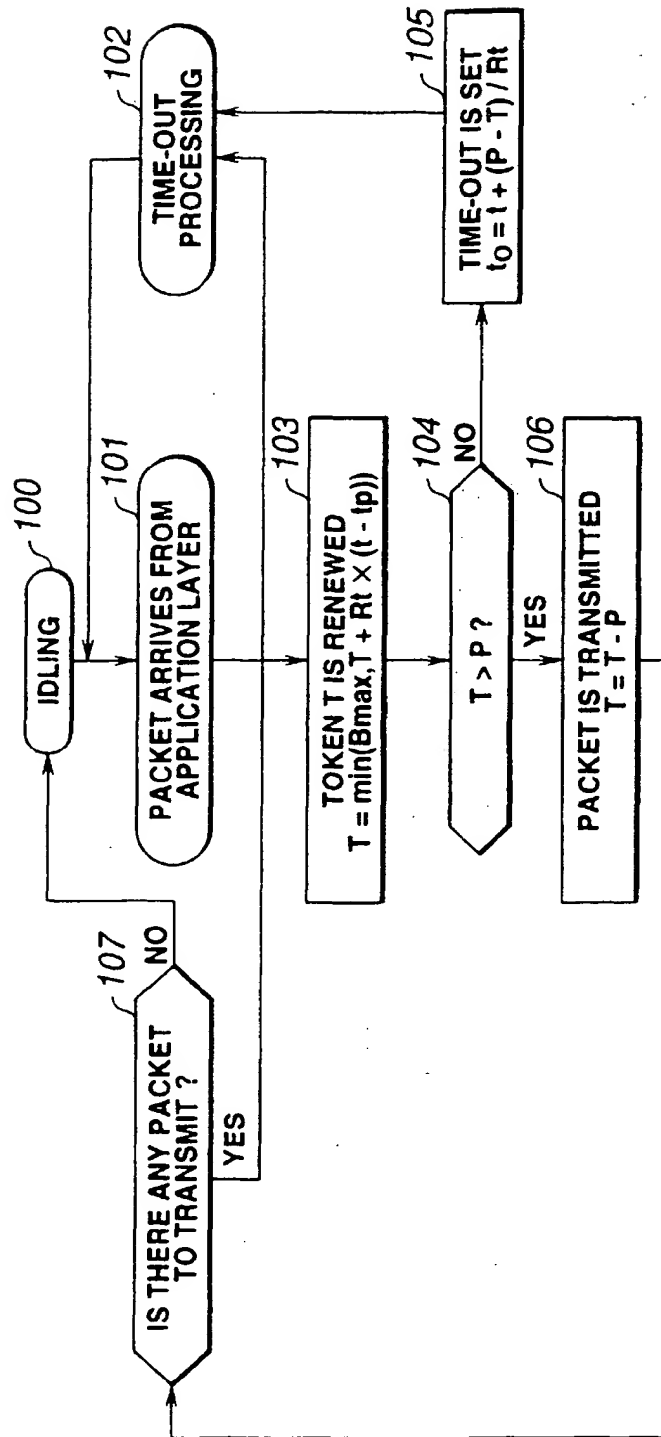
YES

NO

NO

YES

9

European Patent
Office

EUROPEAN SEARCH REPORT

Application Number

EP 98 11 8200

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	GB 2 282 474 A (INTEL CORP) 5 April 1995 * page 9, line 18 - page 10, line 22 * * page 14, line 3 - page 16, line 9 * ---	1,2,4	H04L29/06
Y	EP 0 608 981 A (NIPPON ELECTRIC CO) 3 August 1994 * abstract * ---	1,2,4	
A	US 5 367 523 A (CHANG RONG-FENG ET AL) 22 November 1994 * column 11, line 25 - column 15, line 43 * ---	1,3	
A	POO G -S ET AL: "Modularity versus efficiency in OSI system implementations" COMPUTER COMMUNICATIONS, vol. 20, no. 8, 25 August 1997, page 628-638 XP004126716 * page 633, left-hand column, line 24 - line 46 * -----	1	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6) H04L

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	27 January 1999	RAMIREZ DE AREL.., F

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03 82 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**   EP 98 11 8200

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

27-01-1999

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
GB 2282474	A	05-04-1995	SG	47794 A	17-04-1998
			US	5721955 A	24-02-1998
EP 0608981	A	03-08-1994	JP	2062738 C	24-06-1996
			JP	6209347 A	26-07-1994
			JP	7093645 B	09-10-1995
			CA	2113250 A,C	12-07-1994
			US	5541987 A	30-07-1996
US 5367523	A	22-11-1994	NONE		

EPO FORM P0459